Chapitre 1

Espaces probabilisés

1.1 Introduction

Certains phénomènes observables à notre échelle ont un caractère déterministe, dans le sens où l'on sait prédire avec certitude, ou tout au moins avec une précision acceptable, leur résultat. C'est notamment le cas des phénomèmes régis par les lois de la physique. Ainsi, si l'on connaît la résistivité ρ d'un fil de cuivre, sa longueur l et l'aire s de sa section, l'on peut calculer sa résistance $R = \frac{\rho l}{s}$ puis, connaissant l'intensité I du courant électrique qui le parcourt, prédire la valeur $U = RI$ de la différence de potentiel entre ses deux extrémités. Cette prédiction est déterministe. Si maintenant l'on effectue, à l'aide d'un voltmètre, plusieurs mesures de cette différence de potentiel, l'on constate que le résultat peut varier d'une mesure à l'autre. En effet, la mesure est entachée d'incertitude, du fait de divers facteurs, comme la température, ou la précision et la sensibilité du dispositif de mesure. Ce caractère incertain en fait une variable aléatoire[1], dont on constate néanmoins qu'elle présente de faibles fluctuations[2]

Ainsi, l'on oppose les phénomèmes aléatoires, ou non déterministes, pour lesquels le résultat ne peut être prédit avec certitude, aux phénomènes déterministes dont l'issue est prédictible.

La théorie des probabilités, formalisée par Andreï Nikolaïevitch Kolmogorov, a pour objet la modélisation des phénomènes aléatoires, pour lesquels le hasard intervient. Nous tenterons, dans le présent ouvrage, d'en exposer les fondements, tout en montrant qu'ils reposent en fait sur un certain nombre de constats de simple bon sens.

1.2 Expérience aléatoire — Référentiel

Le premier de ces constats est que, si l'on veut pouvoir modéliser ces phénomènes aléatoires, leur degré d'incertitude ne peut néanmoins être total. On doit *a minima*

[1] Du latin *alea* qui signifie jeu de dés, hasard.
[2] Nous verrons par la suite que l'écart-type est une mesure de ces fluctuations.

être capable d'en décrire l'ensemble des résultats, ou issues, possibles, que ce soit en extension, c'est-à-dire en en dressant la liste, ou bien par compréhension, en énonçant cette fois une propriété commune à toutes ces issues.

Intéressons-nous donc à un phénomène aléatoire, que nous appellerons expérience aléatoire, ou encore épreuve aléatoire[3], et désignerons par la lettre $\mathcal{E}$. Nous pouvons alors donner la

Définition 1 (référentiel) *Soit $\mathcal{E}$ une expérience aléatoire. L'ensemble des résultats possibles de cette expérience est appelé référentiel, ou univers, ou encore espace des réalisations, et noté Ω, tandis que les éléments de cet ensemble, c'est-à-dire les issues possibles de l'expérience, sont désignés par la lettre minuscule ω.*

Exemples

1. `Question :` Soit $\mathcal{E}$ l'expérience aléatoire consistant à lancer un dé puis à noter le chiffre obtenu. Décrire l'espace Ω de ses réalisations.

   Réponse : Il s'agit de l'ensemble fini $\Omega = \{1, 2, 3, 4, 5, 6\}$. Il est ici décrit en extension.

2. `Question :` Soit maintenant $\mathcal{E}$ l'expérience aléatoire consistant à compter le nombre de véhicules se présentant un lundi matin entre 8h et 9h au péage autoroutier de Saint Quentin Fallavier[4]. En distinguant le cas où l'on connaît un majorant N de ce nombre de celui où l'on n'en connaît pas, donner l'univers Ω de cette expérience.

   Réponse : Si l'on connaît un majorant N du nombre de ces véhicules, l'univers ce cette expérience un ensemble fini Ω composé d'entiers compris entre 0 et N. Dans le cas contraire, cet univers est l'ensemble $\Omega = \mathbb{N}$ des entiers naturels. Il est dans ce cas infini dénombrable.

3. `Question :` Si l'on s'intéresse enfin à l'expérience aléatoire $\mathcal{E}$ consistant à mesurer la taille précise d'hommes agés de 20 ans, quel est alors son référentiel Ω ?

   Réponse : Il s'agit cette fois d'un intervalle $\Omega = [a, b]$ de la droite réelle, de sorte que c'est alors un ensemble infini non dénombrable, décrit en compréhension.

1.3 Événements — Tribu d'événements

1.3.1 Événements — Premières conséquences

Donnons tout d'abord la

Définition 2 (événement) *Soit $\mathcal{E}$ une expérience aléatoire de référentiel Ω. Un événement A est dit lié à $\mathcal{E}$ si l'on sait dire, quelle que soit l'issue ω de cette expérience, s'il est ou non réalisé.*

[3]On utilise le terme expérience lorsqu'on maîtrise les conditions nécessaires à l'observation du phénomène, comme par exemple lors d'un essai de laboratoire, et le mot épreuve dans le cas contraire, comme lors de l'observation du temps qu'il fait. Notons toutefois que le terme d'expérience tend de nos jours à supplanter celui d'épreuve.

[4]Département de l'Isère, France métropolitaine.

Remarque Un événement est donc une proposition logique à laquelle l'on sait attribuer, quelle que soit l'issue ω de l'expérience $\mathcal{E}$, la valeur de vérité V (vrai) ou F (faux).

Il découle alors de la définition 2 qu'à tout événement A lié à l'expérience aléatoire $\mathcal{E}$ correspond une partie de son référentiel Ω, que nous désignerons également par A. Inversement, à tout élément A de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω correspond un événement que nous désignerons par la même lettre. Ainsi, l'ensemble des événements liés à $\mathcal{E}$ est équipotent à celui des parties de son référentiel Ω, si bien que de cette équipotence résulte, ainsi que nous allons le voir, une correspondance biunivoque entre les opérations logiques que l'on effectue sur les événements et les opérations ensemblistes sur les parties de Ω.

1.3.2 Logique des événements

Nous pouvons à présent donner la

Définition 3 (logique des événements) *Soit $\mathcal{E}$ une expérience aléatoire de référentiel Ω, et soit ω son issue. Soient par ailleurs A et B des événements liés à $\mathcal{E}$.*

1. **négation non** *Si ω appartient à A, l'événement A est réalisé. Sinon, ω appartient au complémentaire $\overline{A}$ de A dans Ω, et c'est alors l'événement contraire* (non A) *qui est réalisé. Ainsi, à la négation logique est associée l'opération ensembliste de complémentation.*
2. **conjonction et** *L'événement* (A et B) *est réalisé si A est réalisé et si B l'est également, c'est-à-dire si l'issue ω de $\mathcal{E}$ appartient à $A \cap B$, si bien qu'à la conjonction logique correspond l'opération ensembliste d'intersection.*
3. **disjonction ou** *L'événement* (A ou B) *est réalisé si A est réalisé ou si B est réalisé, c'est-à-dire si ω appartient à $A \cup B$, de sorte qu'à la disjonction logique est associée l'opération ensembliste de réunion.*
4. **implication** $\Longrightarrow$ *L'événement A implique l'événement B si sa réalisation entraîne celle de B, ce qui équivaut à dire que $A \subset B$. Ainsi, à l'implication logique correspond la relation ensembliste d'inclusion.*
5. **équivalence** $\Longleftrightarrow$ *Les événements A et B sont équivalents si la réalisation de l'un d'entre eux entraîne celle de l'autre, donc si $A = B$. Il en découle alors qu'à l'équivalence logique est associée la relation ensembliste d'égalité.*
6. **différence** $\setminus$ *L'événement $A \setminus B$ est réalisé si A est réalisé et si B ne l'est pas, ce qui montre qu'à l'opération ensembliste de différence correspond la négation* (non $\Longrightarrow$) *de l'implication logique.*
7. **disjonction exclusive** XOR[5] *L'événement* (A XOR B) *est réalisé si l'un et l'un seulement des événements A et B est réalisé, c'est-à-dire si ω appartient à la différence symétrique $A \Delta B$ de A et B, de sorte qu'à la disjonction logique exclusive est associée l'opération ensembliste de différence symétrique.*

[5] Il n'existe pas de mot français signifiant disjonction exclusive. L'acronyme XOR est la contraction de l'anglais eXclusive OR (ou exclusif).

Le tableau 1.1 regroupe les correspondances entre opérations ou relations logiques et opérations ou relations ensemblistes fournies par la définition 3

TAB. 1.1 – Correspondances entre opérations ou relations logiques et ensemblistes

Opération ou relation logique	Opéation ou relation ensembliste
négation non	complémentation $\overline{}$
conjonction et	intersection $\cap$
disjonction ou	réunion $\cup$
implication $\Longrightarrow$	inclusion $\subset$
équivalence $\Longleftrightarrow$	égalité $=$
négation de l'implication (non $\Longrightarrow$)	différence $\setminus$
disjonction exclusive XOR	différence symétrique Δ

Remarques

1. On a coutume, pour les opérations sur les événements, d'utiliser plutôt les symboles ensemblistes que les symboles logiques. C'est ce que nous ferons dans ce qui suit, en ayant toutefois présent à l'esprit qu'il s'agit de logique des événements.
2. On a
$$A \setminus B = A \cap \overline{B}$$
ainsi que
$$A \Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A) = (A \cap \overline{B}) \cup (B \cap \overline{A})$$
3. On gardera présent à l'esprit qu'en mathématiques, la disjonction logique ou est inclusive. En effet, l'événement $A \cup B$, réalisé lorsque A ou B est réalisé, l'est en particulier lorsqu'ils le sont simultanément. En langue naturelle, le terme « ou » exclut cette dernière possibilité : si le menu d'un restaurant indique « *fromage ou dessert* », il ne faut certainement pas vous attendre à avoir les deux. En revanche, la disjonction exclusive XOR a la même signification que celle du terme « ou » en langue naturelle.
4. L'événement constitué par le référentiel Ω, dont l'issue est certaine, est pour cela appelé événement certain. À l'opposé, son contraire $\emptyset$, qui pour sa part n'advient jamais, est dénommé événement impossible.
5. Les événements que constituent les singletons $\{\omega\}$, où ω parcourt l'univers Ω, et qui ne se réalisent que d'une seule façon, sont qualifiés d'événements élémentaires.
6. Lorsque le conjonction $A \cap B$ des événements A et B se réduit à l'événement impossible $\emptyset$, ce qui signifie que ces deux événement ne peuvent se réaliser simultanément, l'on dit alors qu'ils sont incompatibles.

1.3.3 Tribu d'événements

Lorsque le référentiel Ω d'une expérience aléatoire $\mathcal{E}$ est dénombrable, il n'y a aucune difficulté à considérer l'ensemble $\mathcal{P}(\Omega)$ de tous les événements liés à cette expérience, de sorte que toutes les opérations logiques sur ces derniers fournies par la définition 3 sont alors possibles.

En revanche, si ce référentiel est non dénombrable, comme c'est par exemple le cas si Ω est un intervalle $[a, b]$ de la droite réelle, ou la droite réelle elle même, il devient difficile de considérer l'ensemble de ses parties. Ainsi, l'ensemble $\mathcal{P}(\mathbb{R})$ des parties de $\mathbb{R}$ contient des ensembles non mesurables, qui expliquent notamment le célèbre paradoxe de Banach-Tarski[6], illustré par la figure 1.1.

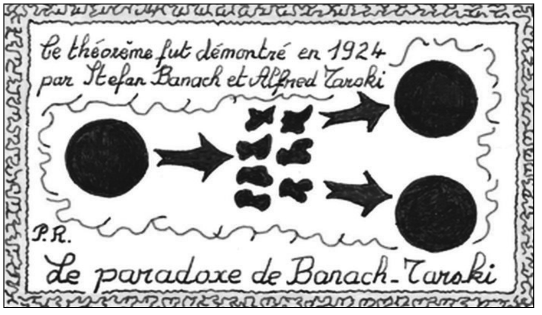

FIG. 1.1 – Illustration du paradoxe de Banach-Tarski

On est donc conduit, dans ce cas, à considérer un sous-ensemble strict de $\mathcal{P}(\Omega)$, plus simple et doté de meilleures propriétés que ce dernier. Il faut toutefois que toutes les opérations logiques sur les événements de ce sous-ensemble données par la définition 3 restent possibles. La structure algébrique la mieux adaptée est alors celle de tribu. C'est la

Définition 4 (tribu d'événements) *Soit $\mathcal{E}$ une espérience aléatoire de référentiel Ω. On appelle tribu sur Ω, ou σ-algèbre, ou, plus rarement, corps de Borel, et l'on désigne par $\mathcal{T}$, toute famille de parties de Ω telle que*

1. *L'événement certain Ω appartient à $\mathcal{T}$.*
2. *Pour tout événement A de $\mathcal{T}$, l'événement contraire $\overline{A}$ appartient à $\mathcal{T}$ (stabilité par complémentation).*
3. *Pour toute famille $(A_i)_{i\in I}$ d'événements de $\mathcal{T}$ indexés par les éléments i de l'ensemble dénombrable I, la disjonction*

$$\bigcup_{i\in I} A_i$$

des ces événements appartient à $\mathcal{T}$ (stabilité par réunion dénombrable).

[6]Ce théorème, démontré en 1924 par Stefan BANACH et Alfred TARSKI, établit qu'il est possible de découper une boule de l'espace eucidien $\mathbb{R}^3$ en un nombre fini de morceaux, puis de réassembler ces derniers sans les déformer, de façon à obtenir deux boules identiques à la première.

Exemples

1. L'ensemble $\mathcal{P}(\Omega)$ des parties de Ω est une tribu sur Ω, appelée tribu discrète, ou tribu la plus fine.
2. L'ensemble $\{\Omega, \emptyset\}$ est une tribu sur Ω, appelée tribu grossière.
3. Si A est une partie de Ω, l'ensemble $\{\Omega, \emptyset, A, \overline{A}\}$ est une tribu sur Ω. Nous verrons dans ce qui suit qu'elle n'est autre que la tribu de Bernoulli[7] engendrée par A.

Remarques

1. Si $\mathcal{T}$ est une tribu sur Ω, le couple $(\Omega, \mathcal{T})$ est dénommé espace probabilisable.
2. Dans tout ce qui suit, seuls les éléments de la tribu $\mathcal{T}$ d'un espace probabilisable $(\Omega, \mathcal{T})$ seront dénommés événements.

Les trois propriétés d'une tribu d'événements fournies par la définition 4 suffisent alors pour garantir la stabilité de cette tribu pour les autres opérations logiques de la définition 3, comme l'illustre la preuve du

Théorème 1 (stabilité d'une tribu d'événements) *Soit $\mathcal{E}$ une espérience aléatoire de référentiel Ω et soit $\mathcal{T}$ une tribu sur Ω. Alors*

1. *L'événement impossible $\emptyset$ appartient à $\mathcal{T}$.*
2. *Pour toute famille $(A_i)_{i \in I}$ d'événements de $\mathcal{T}$ indexés par les éléments i de l'ensemble dénombrable I, la conjonction*

$$\bigcap_{i \in I} A_i$$

   de ces événements appartient à $\mathcal{T}$.

3. *Pour tous événements A et B de $\mathcal{T}$, la différence $A \setminus B$ de B dans A ainsi que la différence symétrique $A \Delta B$ de ces événements appartient à $\mathcal{T}$.*

Preuve

1. Comme l'événement impossible $\emptyset$ est le complémentaire de l'événement certain Ω, il découle de la deuxième propriété d'une tribu qu'il appartient également à $\mathcal{T}$.
2. Les événements $\overline{A_i}$, $i \in I$, appartiennent à la tribu $\mathcal{T}$ en vertu de la deuxième propriété, et il en est de même de leur disjontion, compte tenu cette fois de la troisième propriété, ainsi que du complémentaire de cette disjonction, grâce à nouveau à la deuxième propriété. Or l'on a

$$\overline{\bigcup_{i \in I} \overline{A_i}} = \bigcap_{i \in I} \overline{\overline{A_i}} = \bigcap_{i \in I} A_i$$

   ce qui montre que la conjonction des événements A_i, $i \in I$, appartient aussi à la tribu $\mathcal{T}$.

[7] Jacob Bernoulli (1654-1705) est un mathématicien et physicien suisse.

3. Enfin, si A et B sont deux événements de $\mathcal{T}$, alors, $\overline{A}$ et $\overline{B}$ appartiennent aussi à $\mathcal{T}$ en vertu de la deuxième propriété. Il en est donc de même, compte tenu du point précédent, de la conjonction $A \cap \overline{B}$, qui n'est autre que la différence de B dans A, ainsi que de la conjonction $\overline{A} \cap B$, et donc, compte tenu cette fois de la troisième propriété, de leur disjonction, $(A \cap \overline{B}) \cup (\overline{A} \cap B)$, c'est-à-dire de la différence symétrique des événements A et B, ce qui achève la preuve du théorème 1. □

Établissons à présent le

Théorème 2 (intersection de tribus) *Soit $\mathcal{E}$ une espérience aléatoire de référentiel Ω, et soit $(\mathcal{T}_i)_{i \in I}$, une famille de tribus sur Ω, indexées par les éléments i de l'ensemble dénombrable I. Alors, leur intersection*

$$\mathcal{T} = \bigcap_{i \in I} \mathcal{T}_i$$

est une tribu sur Ω.

Preuve Comme Ω appartient à chacune des tribus $\mathcal{T}_i$, $i \in I$, il appartient aussi à leur intersection $\mathcal{T}$. Si maintenant A est un événement de $\mathcal{T}$, il appartient à chacune des tribus $\mathcal{T}_i$, $i \in I$, et il en est donc de même de son contraire $\overline{A}$, qui appartient alors à $\mathcal{T}$. Si enfin $(A_j)_{i \in J}$ est une famille dénombrable d'événements de $\mathcal{T}$, c'est une famille d'événements de chacune des tribus $\mathcal{T}_i$, $i \in I$, de sorte que la conjonction de ces événements appartient aussi à chacune des tribus $\mathcal{T}_i$, $i \in I$, et donc à leur intersection $\mathcal{T}$, qui satisfait ainsi les trois propriétés d'une tribu, ce qui achève la preuve du théorème 2. □

Pour construire une tribu $\mathcal{T}$ sur Ω, l'on part généralement d'une famille $\mathcal{F}$ réduite de parties de Ω. Cette construction repose sur la

Définition 5 (tribu engendrée par une famille d'événements) *Soit $\mathcal{E}$ une espérience aléatoire de référentiel Ω, et soit $\mathcal{F}$ une famille de parties de Ω. La tribu $\mathcal{T}$ engendrée par $\mathcal{F}$, que l'on note également $\langle \mathcal{F} \rangle$, est la plus petite tribu contenant $\mathcal{F}$ au sens de l'inclusion dans $\mathcal{P}(\Omega)$.*

Remarque Si $\mathcal{F}$ est une famille de parties de Ω, l'existence de la tribu $\langle \mathcal{F} \rangle$ qu'elle engendre est assurée, puisque $\mathcal{P}(\Omega)$ est une tribu sur Ω contenant $\mathcal{F}$.

La tribu $\mathcal{T}$ engendrée par une famille $\mathcal{F}$ d'événements est caractérisée par le

Théorème 3 (tribu engendrée par une famille d'événements) *Soit $\mathcal{E}$ une espérience aléatoire de référentiel Ω, et soit $\mathcal{F}$ une famille de parties de Ω. La tribu $\mathcal{T}$ engendrée par $\mathcal{F}$ est l'intersection de toutes les tribus contenant $\mathcal{F}$.*

Preuve Soit $(\mathcal{T}_i)_{i \in I}$ la famille de toutes les tribus sur Ω contenant $\mathcal{F}$, et soit $\mathcal{T}$ leur intersection. Alors, d'après le théorème 2, $\mathcal{T}$ est une tribu sur Ω contenant $\mathcal{F}$ et incluse dans toute tribu contenant $\mathcal{F}$. C'est donc la plus petite tribu contenant $\mathcal{F}$ pour la relation d'inclusion dans $\mathcal{P}(\Omega)$, ce qui prouve le théorème 3. □

Exemples

1. `Question` : Soit A est une partie de Ω. Décrire la tribu qu'elle engendre.
   Réponse : Il s'agit la tribu de Bernoulli $\mathcal{T} = \{\Omega, \emptyset, A, \overline{A}\}$.
2. `Question` : Si A et B sont deux événements incompatibles, quelle est la tribu qu'ils engendrent ?
   Réponse : Il s'agit cette fois de la tribu $\mathcal{T} = \{\Omega, \emptyset, A, \overline{A}, B, \overline{B}, A \cup B, \overline{A} \cap \overline{B}\}$.

Donnons enfin la

Définition 6 *Soit $(\Omega_i, \mathcal{T}_i)_{i \in I}$ une famille d'espaces probabilisables indexés par les éléments i de l'ensemble dénombrable I, et soit Ω le produit cartésien des ensembles Ω_i, $i \in I$. On appelle tribu produit sur Ω, et l'on note*

$$\mathcal{T} = \bigotimes_{i \in I} \mathcal{T}_i$$

la tribu engendrée par la famille $\mathcal{F}$ de tous les événements de Ω contitués par les produits cartésiens d'événements A_i, $i \in I$, où, pour chaque indice i, A_i parcourt la tribu $\mathcal{T}_i$.

1.4 Probabilité sur un espace probabilisable

1.4.1 Définition

Soit $\mathcal{E}$ une expérience aléatoire de référentiel Ω, et soit $\mathcal{T}$ une tribu sur Ω. Il s'agit à présent de munir l'espace probabilisable $(\Omega, \mathcal{T})$ d'une mesure permettant de quantifier les chances de réalisation des événements de la tribu $\mathcal{T}$. Cette mesure, dénommée probabilité sur l'espace probabilisable $(\Omega, \mathcal{T})$, et notée P, consiste à affecter à chaque événement A de la tribu $\mathcal{T}$ un nombre réel, compris entre 0 et 1, et noté $P(A)$, quantifiant les chances de réalisation de cet événement. Elle doit alors posséder certaines propriétés découlant de constats de bon sens. Tout d'abord, plus un événement a de chances de se réaliser, plus sa probabilité doit être élevée. Ainsi, l'événement certain Ω, dont l'issue est acquise, doit avoir la probabilité la plus élevée, soit 1. À l'opposé, l'événement impossible $\emptyset$, qui n'advient jamais, doit posséder la probabilité la plus faible, c'est-à-dire 0. Par ailleurs, si l'événement A implique l'événement B, ses chances de réalisation doivent être plus faibles que celles de ce dernier. Enfin, si les événements A et B sont incompatibles, la probabilité de leur disjonction $A \cup B$ doit être la somme de leurs probabilités respectives. Ces propriétés conduisent alors à la

Définition 7 (probabilité sur un espace probabilisable) *Soit $\mathcal{E}$ une expérience aléatoire de référentiel Ω, et soit $\mathcal{T}$ une tribu sur Ω. On appelle probabilité sur l'espace probabilisable $(\Omega, \mathcal{T})$ toute application P de $\mathcal{T}$ dans $[0, 1]$ telle que*

1. *La probabilité de l'événement certain Ω est égale à 1.*
2. *Pour toute famille $(A_i)_{i \in I}$ d'événements de $\mathcal{T}$ indexés par les éléments i de l'ensemble dénombrable I, deux à deux incompatibles, l'on a (σ-additivité)*

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$$

Le triplet $(\Omega, \mathcal{T}, P)$ est alors appelé espace probabilisé.